AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application:

Claim underlining shows the changes from the originally issued patent.

Cancel Claims 1-26.

1-26. (Cancelled)

Add Claim 27 as follows:

27. A method of parallelizing an operation, the method comprising the steps of:

dividing the operation into a set of work partitions;

assigning work partitions from said set of work partitions to a plurality of entities,

wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;

wherein the step of assigning work partitions is performed by assigning the work partitions in a sequence based at least in part on sizes associated with the work partitions, with relatively larger work partitions assigned before relatively smaller work partitions;

said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said operation; and

wherein assigning the work partitions in a sequence includes assigning a first previously unassigned work partition to a particular entity of the plurality of entities, and when the particular entity completes processing the first work partition, picking a second previously unassigned work partition based at least in part to the size of the second work partition, and assigning the second unassigned work partition to the particular entity for processing,

wherein the method is performed by one or more computing devices.

Cancel Claim 28.

28. (Canceled)

Add Claims 29-52 as follows:

29. A method of parallelizing an operation, the method comprising the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions, wherein the step of assigning work partitions includes:
assigning said at least one entity a first work partition from said set of work partitions; and
after said at least one entity has completed operation on said first work partition,
assigning said at least one entity a second work partition from said set of work partitions, wherein the step of assigning said at least one entity a second work partition includes
determining whether there are any unassigned work partitions from a first level in a hierarchy to which said first work partition belonged; and
if there are no unassigned work partitions from the first level in the hierarchy, then selecting said second work partition from a level in said hierarchy that is two levels above said first level in said hierarchy;
said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said operation; and
wherein the operation is specified in a query that corresponds to the hierarchy of operations,
wherein the method is performed by one or more computing devices.

30. A method of parallelizing an operation, the method comprising the steps of:
dividing the operation into a set of work partitions;

assigning work partitions from said set of work partitions to a plurality of entities, wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;

said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said operation;

the method includes the step of generating a serial execution plan for operations in a database management system (DBMS) running on a computer system;

the method includes the step of generating a parallelized execution plan for said serial execution plan, said parallelized execution plan including first and second operations;

the step of dividing an operation is performed by dividing said second operation;

the plurality of entities includes one or more slave processes operating on a plurality of data partitions, the quantity of said data partitions being greater than the quantity of said slave processes;

executing said parallelized execution plan when a plurality of parallel resources of said computer system are available; and

executing said serial execution plan when said plurality of resources are not available, wherein the method is performed by one or more computing devices.

31. The method of claim 30 wherein said step of generating a parallelized execution plan includes the steps of:

identifying one or more segments of said serial execution plan that can be parallelized; and

identifying partitioning requirements of said one or more segments.

32. The method of claim 30 wherein said step of generating a parallelized execution plan is based on a specification of parallelism in a statement specifying one of said operations.

33. A method of parallelizing an operation, the method comprising the steps of:

dividing the operation into a set of work partitions;

assigning work partitions from said set of work partitions to a plurality of entities, wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;

said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said operation;

generating an execution plan for said operation;

examining said execution plan from bottom up;

identifying a parallelized portion of said execution plan, said parallelized portion can be processed in parallel, said parallelized portion including first and second operations, said first and second operations being executable in parallel;

wherein the step of dividing the operation is performed by dividing said second operation;

wherein the plurality of entities includes one or more slave processes operating on a plurality of data partitions, the quantity of said data partitions being greater than the quantity of said slave processes;

identifying some serial portion of said execution plan, said serial portion can be processed in serial; and

allocating a central scheduler between said parallelized portion and said serial portion, wherein the method is performed by one or more computing devices.

34. The method of Claim 33 further including the steps of:

identifying a first data flow requirement for a first portion of said execution plan said first data flow requirement corresponding to a partitioning of a data flow required by said first portion;

identifying a second data flow requirement for a second portion of said execution plan said second data flow requirement corresponding by said second portion; and

allocating a data flow director between said first portion and said second portion when said first data flow requirement is not compatible with said second data flow requirement said data flow director repartitioning a data flow of said first portion to be compatible with said second data flow requirement.

35. A method for parallelizing an operation, the method comprising the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities, wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;
said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said operation;
generating an execution plan to execute database management system (DBMS) operations in parallel, said execution plan including first and second operations;
wherein the step of dividing said operation is performed by dividing said second operation;
initiating an operation coordinator in a computer system to coordinate execution of said execution plan;
initiating, by said operation coordinator, a first set of slaves operating on a plurality of data partitions to produce data, the quantity of said data partitions being greater than the quantity of said first set of slave processes;
initiating, as said plurality of entities, by said operation coordinator, a second set of slaves to consume data; and
directing said second set of slaves to produce data and said first set of slaves to consume data when said first set of slaves finishes producing data,
wherein the method is performed by one or more computing devices.

36. The method of claim 35 wherein said execution plan is comprised of operator nodes and said operator nodes are linked together to form execution sets.

37. A method for parallelizing an operation, the method comprising the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities, wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;

said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said operation;

generating an execution plan to execute said operations in parallel, said execution plan including first and second operations;

wherein the step of dividing said operation includes dividing said first operation;

initiating producer slaves operating on a plurality of data partitions to produce a first data production;

initiating consumer slaves to consume said first data production;

when said first data production is completed, generating an identification of a plurality of said consumer slaves that did not receive data in said first data production;

examining said identification during a subsequent data production; and

reducing said subsequent data production such that said subsequent data production does not produce data for said plurality of said consumer slaves,

wherein the method is performed by one or more computing devices.

38. A method for processing a statement in a database system, the method comprising the steps of:

receiving, at a database server, a statement that specifies at least a database operation that operates on data within a database;

determining, at said database server, a user-specified degree of parallelism to use in performing the database operation, wherein said user-specified degree of parallelism expressly indicates a specific number of entities to use in parallel to perform said database operation;

dividing, at said database server, the database operation into a set of work partitions;

performing, at said database server, a determination of how many entities to use to perform said operation based, at least in part, on the user-specified degree of parallelism, wherein the amount of entities that are chosen to use to perform on the database operation is different than the amount of entities that would have been chosen if no user-specified degree of parallelism had been specified;

assigning, at said database server, work partitions from said set of work partitions to a plurality of entities based on said determination; and

said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said database operation,

wherein the method is performed by one or more computing devices.

39. The method of Claim 38 wherein:

the statement requires a plurality of operations;

the user-specified degree of parallelism is specified in said statement, and

the statement specifies said degree of parallelism for a subset of the plurality of operations required by the statement.

40. The method of Claim 38 wherein

the user-specified degree of parallelism is specified in said statement; and

the degree of parallelism specified by the statement indicates that no amount of parallelism is to be used during execution of a particular portion of the statement.

41. The method of Claim 38 wherein

the user-specified degree of parallelism is specified in said statement, and

the degree of parallelism specified by the statement indicates a maximum amount of parallelism to use during execution of said operation.

42. A method of processing a query in a database system, the method comprising the steps of:

dividing, at a database server, a database operation required by said query into a set of work partitions by generating a set of query fragments, each work partition of said set of work partitions to be performed serially by a single entity to which said work partition is assigned;

incorporating hints into at least some of said query fragments at said database server, wherein said query fragments incorporating hints comprise work partitions that may be performed in a plurality of ways to reach a same result, and wherein said

hint associated with a given query fragment indicates one way of said plurality of ways to perform said work partition;

assigning, at said database server, query fragments from said set of query fragments to a plurality of entities; and

said plurality of entities operating in parallel on query fragments assigned to said plurality of entities to perform said database operation, wherein entities working on a query fragment associated with a hint perform the work partition associated with said query fragment in said one way dictated by said hint,

wherein the method is performed by one or more computing devices.

43. The method of Claim 42 wherein the step of incorporating hints includes incorporating hints that dictate the operation of a table scan.

44. The method of Claim 43 wherein the step of incorporating hints that dictate the operation of a table scan includes incorporating hints that rowid partitioning is to be used during the table scan.

45. The method of Claim 42 wherein the step of incorporating hints includes incorporating hints that specify performance of a full table scan.

46. The method of Claim 42 wherein the step of incorporating hints includes incorporating hints that specify using a particular type of join.

47. The method of Claim 46 wherein the step of incorporating hints that specify using a particular type of join includes incorporating hints that specify using a sort/merge join.

48. The method of Claim 46 wherein the step of incorporating hints that specify using a particular type of join includes incorporating hints that specify using a nested loop join.

49. A method of processing a query, the method comprising the steps of:

determining a hierarchy of operations associated with a query;

dividing a first operation required by said query into a first set of work partitions;

dividing a second operation required by said query into a second set of work partitions, wherein said second operation immediately follows said first operation in said hierarchy;

dividing a third operation required by said query into a third set of work partitions, wherein said third operation immediately follows said second operation in said hierarchy;

assigning work partitions from said first set of work partitions to a first plurality of entities;

said first plurality of entities operating in parallel on work partitions assigned to said first plurality of entities from said first set of work partitions to perform said first operation;

assigning work partitions from said second set of work partitions to a second plurality of entities, wherein said second plurality of entities are different entities than said first plurality of entities; and

said second plurality of entities operating in parallel on work partitions assigned to said second plurality of entities from said second set of work partitions to perform said second operation;

assigning work partitions from said third set of work partitions to said first plurality of entities; and

said first plurality of entities operating in parallel on work partitions assigned to said first plurality of entities from said third set of work partitions to perform said third operation,

wherein the method is performed by one or more computing devices.

50. The method of Claim 49 further comprising performing the following steps when a given entity in said first set of entities finishes performing a work partition from said first set of work partitions:

determining whether there are any unassigned work partitions from said first set of work partitions; and

if there are no unassigned work partitions from said first set of work partitions, then assigning the given entity a work partition selected from said third set of work partitions; and

if there are unassigned work partitions from said first set of work partitions, then assigning the given entity a work partition selected from said first set of work partitions.

51. The method of Claim 49 wherein the hierarchy includes odd levels and even levels, and the method further comprises the steps of assigning work partitions from odd levels to said first plurality of entities and work partitions from even levels to said second plurality of entities.

52. The method of Claim 49 wherein performing work partitions in said first set of work partitions causes said first set of entities produce output consumed by said second plurality of entities, and performing work partitions in said third set of work partitions causes said first set of entities to consume output produced by said second plurality of entities.

Cancel Claims 53-62.

53-62. (Cancelled)

Add Claim 63 as follows:

63. A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:

dividing the operation into a set of work partitions;

assigning work partitions from said set of work partitions to a plurality of entities, wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;

wherein the step of assigning work partitions is performed by assigning the work partitions in a sequence based at least in part on sizes associated, with the work partitions with relatively larger work partitions assigned before relatively smaller work partitions;

said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said operation; and

wherein assigning the work partitions in a sequence includes assigning a first previously unassigned work partition to a particular entity of the plurality of entities, and when the particular entity completes processing the first work partition, picking a second previously unassigned work partition based at least in part to the size of the second work partition, and assigning the second unassigned work partition to the particular entity for processing.

Cancel Claim 64.

64. (Canceled)

Add Claims 65-88 as follows:

65. A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:

dividing the operation into a set of work partitions;

assigning work partitions from said set of work partitions to a plurality of entities, wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions, wherein the step of assigning work partitions includes

assigning said at least one entity a first work partition from said set of work partitions; and

after said at least one entity has completed operating on said first work partition,
assigning said at least one entity a second work partition from said set of work partitions;
said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said operation;
wherein the operation is specified in a query that corresponds to a hierarchy of operations; and
the step of assigning said at least one entity a second work partition includes
determining whether there are any unassigned work partitions from a first level in the hierarchy to which said first work partition belonged; and
if there are no unassigned work partitions from the first level in the hierarchy, then selecting said second work partition from a level in said hierarchy that is two levels above said first level in said hierarchy.

66. A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;
said plurality of entities operation in parallel on work partitions assigned to said plurality of entities to perform said operation;
wherein the instructions include instructions for performing the step of generating a serial execution plan for operations in a database management system (DBMS) running on a computer system;
wherein the instructions include instructions for performing the step of generating a parallelized execution plan for said serial execution plan, said parallelized execution plan including first and second operations;
wherein the step of dividing an operation is performed by dividing said second operation;

wherein the plurality of entities includes one or more slave processes operating on a plurality of data partitions, the quantity of said data partitions being greater than the quantity of said slave processes;

wherein the instructions include instructions for performing the step of executing said parallelized execution plan when a plurality of parallel resources of said computer system are available; and

wherein the instructions include instructions for performing the step of executing said serial execution plan when said plurality of resources are not available.

67. The computer-readable storage medium of claim 66 wherein said step of generating a parallelized execution plan includes the steps of:

identifying one or more segments of said serial execution plan that can be parallelized; and

identifying partitioning requirements of said one or more segments.

68. The computer-readable storage medium of claim 66 wherein said step of generating a parallelized execution plan is based on a specification of parallelism in a statement specifying one of said operations.

69. A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:

dividing the operation into a set of work partitions;

assigning work partitions from said set of work partitions to a plurality of entities, wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;

said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform some operation;

generating an execution plan for said operation;

examining said execution plan from bottom up;

identifying a parallelized portion of said execution plan, said parallelized portion can be processed in parallel, said parallelized portion including first and second operations, said first and second operations being executable in parallel;

wherein the step of dividing the operation is performed by dividing said second operation;

wherein the plurality of entities includes one or more slave processes operating on a plurality of data partitions, the quantity of said data partitions being greater than the quantity of said slave processes;

identifying some serial portion of said execution plan, said serial portion can be processed in serial; and

allocating a central scheduler between said parallelized portion and said serial portion.

70. The computer-readable storage medium of Claim 69 further including instructions for performing the steps of:

identifying a first data flow requirement for a first portion of said execution plan said first data flow requirement corresponding to a partitioning of a data flow required by said first portion;

identifying a second data flow requirement for a second portion of said execution plan said second data flow requirement corresponding by said second portion; and

allocating a data flow director between said first portion and said second portion when said first data flow requirement is not compatible with said second data flow requirement said data flow director repartitioning a data flow of said first portion to be compatible with said second data flow requirement.

71. A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:

dividing the operation into a set of work partitions;

assigning work partitions from said set of work partitions to a plurality of entities, wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;

said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said operation;

generating an execution plan to execute database management system (DBMS) operations in parallel, said execution plan including first and second operations;

wherein the step of dividing said operation is performed by dividing said second operation;

initiating an operation coordinator in a computer system to coordinate execution of said execution plan;

initiating, by said operation coordinator, a first set of slaves operating on a plurality of data partitions to produce data, the quantity of said data partitions being greater than the quantity of said first set of slave processes;

initiating, as said plurality of entities, by said operation coordinator, a second set of slaves to consume data; and

directing said second set of slaves to produce data and said first set of slaves to consume data when said first set of slaves finishes producing data.

72. The computer-readable storage medium of claim 71 wherein said execution plan is comprised of operator nodes and said operator nodes are linked together to form execution sets.

73. A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:

dividing the operation into a set of work partitions;

assigning work partitions from said set of work partitions to a plurality of entities, wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;

said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said operation;

generating an execution plan to execute said operations in parallel, said execution plan including first and second operations;

wherein the step of dividing said operation includes dividing said first operation;
initiating producer slaves operating on a plurality of data partitions to produce a first data production;
initiating consumer slaves to consume said first data production;
when said first data production is completed, generating an identification of a plurality of said consumer slaves that did not receive data in said first data production;
examining said identification during a subsequent data production; and
reducing said subsequent data production such that said subsequent data production does not produce data for said plurality of said consumer slaves.

74. A computer-readable storage medium storing instructions for processing a statement in a database system, the instructions including instructions for performing the steps of:
receiving, at a database server, a statement that specifies at least a database operation that operates on data within a database;
determining, at said database server, a user-specified degree of parallelism to use in performing the database operation, wherein said user-specified degree of parallelism expressly indicates a specific number of entities to use in parallel to perform said database operation;
dividing, at said database server, the database operation into a set of work partitions;
performing, at said database server, a determination of how many entities to use to perform said operation based, at least in part, on the user-specified degree of parallelism, wherein the amount of entities that are chosen to use to perform on the database operation is different than the amount of entities that would have been chosen if no user-specified degree of parallelism had been specified;
assigning, at said database server, work partitions from said set of work partitions to a plurality of entities based on said determination; and
said plurality of entities operating in parallel on work partitions assigned to said plurality of entities to perform said database operation,
wherein the method is performed by one or more computing devices.

75. The computer-readable storage medium of Claim 74 wherein:
the statement requires a plurality of operations;
the user-specified degree of parallelism is specified in said statement, and
the statement specifies said degree of parallelism for a subset of the plurality of operations required by the statement.

76. The computer-readable storage medium of Claim 74 wherein
the user-specified degree of parallelism is specified in said statement; and
the degree of parallelism specified by the statement indicates that no amount of parallelism is to be used during execution of a particular portion of the statement.

77. The computer-readable storage medium of Claim 74 wherein
the user-specified degree of parallelism is specified in said statement, and
the degree of parallelism specified by the statement indicates a maximum amount of parallelism to use during execution of said operation.

78. A computer-readable storage medium carrying instructions for processing a query in a database system, the instructions including instructions for performing the steps of:
dividing, at a database server, a database operation required by said query into a set of work partitions by generating a set of query fragments, each work partition of said set of work partitions to be performed serially by a single entity to which said work partition is assigned;
incorporating hints into at least some of said query fragments at said database server, wherein said query fragments incorporating hints comprise work partitions that may be performed in a plurality of ways to reach a same result, and wherein said hint associated with a given query fragment indicates one way of said plurality of ways to perform said work partition;
assigning, at said database server, query fragments from said set of query fragments to a plurality of entities; and

said plurality of entities operating in parallel on query fragments assigned to said plurality of entities to perform said database operation, wherein entities working on a query fragment associated with a hint perform the work partition associated with said query fragment in said one way dictated by said hint,

wherein the method is performed by one or more computing devices.

79. The computer-readable storage medium of Claim 78 wherein the step of incorporating hints includes incorporating hints that dictate the operation of a table scan.

80. The computer-readable storage medium of Claim 79 wherein the step of incorporating hints that dictate the operation of a table scan includes incorporating hints that rowid partitioning is to be used during the table scan.

81. The computer-readable storage medium of Claim 78 wherein the step of incorporating hints includes incorporating hints that specify performance of a full table scan.

82. The computer-readable storage medium of Claim 78 wherein the step of incorporating hints includes incorporating hints that specify using a particular type of join.

83. The computer-readable storage medium of Claim 82 wherein the step of incorporating hints that specify using a particular type of join includes incorporating hints that specify using a sort/merge join.

84. The computer-readable storage medium of Claim 82 wherein the step of incorporating hints that specify using a particular type of join includes incorporating hints that specify using a nested loop join.

85. A computer-readable storage medium carrying instructions for processing a query, the instructions including instructions for performing the steps of:

determining a hierarchy of operations associated with a query;

dividing a first operation required by said query into a first set of work partitions;

dividing a second operation required by said query into a second set of work partitions, wherein said second operation immediately follows said first operation in said hierarchy;

dividing a third operation required by said query into a third set of work partitions, wherein said third operation immediately follows said second operation in said hierarchy;

assigning work partitions from said first set of work partitions to a first plurality of entities;

said first plurality of entities operating in parallel on work partitions assigned to said first plurality of entities from said first set of work partitions to perform said first operation;

assigning work partitions from said second set of work partitions to a second plurality of entities, wherein said second plurality of entities are different entities than said first plurality of entities; and

said second plurality of entities operating in parallel on work partitions assigned to said second plurality of entities from said second set of work partitions to perform said second operation;

assigning work partitions from said third set of work partitions to said first plurality of entities; and

said first plurality of entities operating in parallel on work partitions assigned to said first plurality of entities from said third set of work partitions to perform said third operation.

86. The computer-readable storage medium of Claim 85 further comprising instructions for performing the following steps when a given entity in said first set of entities finishes performing a work partition from said first set of work partitions:

determining whether there are any unassigned work partitions from said first set of work partitions; and

if there are no unassigned work partitions from said first set of work partitions, then assigning the given entity a work partition selected from said third set of work partitions; and

if there are unassigned work partitions from said first set of work partitions, then assigning the given entity a work partition selected from said first set of work partitions.

87. The computer-readable storage medium of Claim 85 wherein the hierarchy includes odd levels and even levels, and the instructions further include instructions for performing the steps of assigning work partitions from odd levels to said first plurality of entities and work partitions from even levels to said second plurality of entities.

88. The computer-readable storage medium of Claim 85 wherein performing work partitions in said first set of work partitions causes said first set of entities produce output consumed by said second plurality of entities, and performing work partitions in said third set of work partitions causes said first set of entities to consume output produced by said second plurality of entities.

Cancel Claims 89-91.

89-91. (Canceled).

Add Claims 92-95 as follows:

92. The method of Claim 38, wherein the user-specified degree of parallelism is specified in said statement.

93. The method of Claim 38, wherein the user-specified degree of parallelism is specified for operations that involve a particular table.

94. The computer-readable storage medium of Claim 74, wherein the user-specified degree of parallelism is specified in said statement.

95. The computer-readable storage medium of Claim 74, wherein the user-specified degree of parallelism is specified for operations that involve a particular table.